working paper
department
of economics

Reputation, Unobserved Strategies,
and Active Supermartingales

by
Drew Fudenberg
and
David K. Levine

Number 490

March 1988

massachusetts
institute of
technology

50 memorial drive
cambridge, mass. 02139

Reputation, Unobserved Strategies, and Active Supermartingales

by
Drew Fudenberg
and
David K. Levine

Number 490 March 1988

REPUTATION, UNOBSERVED STRATEGIES,

AND ACTIVE SUPERMARTINGALES

Drew Fudenberg, Massachusetts Institute of Technology
David K. Levine, University of California, Los Angeles

March 1988

This paper benefited from conversations with David Kreps, John McCall, John Riley, and especially Shmuel Zamir, who suggested that our results would extend to games with moral hazard. The National Science Foundation grants SES 87-08616, SES 86-09697, SES 85-09484, the Sloan Foundation, and the UCLA Academic Senate provided financial support.

1. INTRODUCTION

In this paper we consider a game in which a single long-run player faces a sequence of short-run opponents, each of whom plays only once, but is informed of the outcomes of play in each previous period. These outcomes may not reveal the long-run player's past choices, either because the long-run player's action is subject to moral hazard, or because the long-run player has chosen to play a mixed strategy: In either case, the observed outcomes give only imperfect, probabilistic information about the long-run player's choices. We further assume that the short-run players are uncertain of the long-run player's payoff function, and model this uncertainty by with a probability distribution over the "types" of the long-run player. We focus on "commitment types" who play the same stage-game strategy in every period of play. Our main result is that the long-run player's payoff in any Nash equilibrium is bounded below by an amount that converges, as the discount factor tends to one, to the most he could get by commiting himself to any of the strategies for which the corresponding commitment type has positive probability. A loose way of saying this is that the long-run player can obtain a reputation for always playing any strategy which the short-run players believe has positive probability of always being played. Note that this reputation, and the corresponding lower bound on the long-run player's payoff, depend only on the type that the long-run player prefers to mimic and is independent of the other types that have positive probability. In Fudenberg-Levine [1987] we proved a similar but more restrictive theorem. There, we assumed that the short-run players observe the actions that the long-run player has chosen, and also restricted attention to reputations for playing pure strategies. Under these assumptions, if the long-run player fails to play a strategy in any period the short-run players are certain to learn that he is not the corresponding commitment type. Conversely, if the long-run player plays strategy s in a period where the short-run players do not expect him to do so, the short-run

players are certain to be "surprised." When the short-run players do not directly observe the long-run player's choice of action in the stage game, or when commitment strategies are mixed instead of pure, the short-run players are not certain to detect deviations and our previous analysis does not apply.

One implication of our results is that the long-run player can build a reputation for playing any mixed strategy for which the short-run players assign positive prior probability to the corresponding type. The case of mixed strategies in games without moral hazard is of particular interest in light of the results of Fudenberg-Kreps-Maskin [1987] on repeated games with long-run and short-run players. Fudenberg-Kreps-Maskin showed that that the pure-strategy commitment payoff is not always a tight lower bound on what the long-run player can obtain in any equilibrium, because in some games by playing a mixed strategy the long-run player can induce the short-run players to choose a more favorable response. However, when the long-run player's payoff function is common knowledge, in general he cannot do as well as he could by committing himself to a mixed strategy. Our results here show if the corresponding commitment types have positive prior probability the the long-run player <u>can</u> in fact build a reputation for playing a mixed strategy, and thus attain a higher payoff than in any equilibria of the unperturbed game.

We prove our result as follows: Let γ_1^* denote the distribution over outcomes that corresponds to strategy σ_1^*, and imagine that the short-run players assign positive prior probability to the long-run player being a type that always plays σ_1^*. Since the short-run players are myopic, they will play a best response to γ_1^* in any period where they expect the distribution over outcomes to be close to γ_1^*. Now imagine that the long-run player chooses to always play σ_1^*. In any period where the short-run players do not play a best response to γ_1^*, there is a non-negligible probability that they will revise their posterior beliefs a non-negligible amount in the direction of the

long-run player being a type who always plays σ_1^*. Intuitively, if the short-run players do not play a best response to σ_1^*, they will, with some probability, be "surprised" when σ_1^* is played.

After sufficiently many of these surprises, the short-run players will attach a very high probability to the long-run player playing σ_1^* for the rest of the game, and thus will play best responses to γ_1^* from then on. Thus one would expect that for any ϵ there is an $K(\epsilon)$ such that with probability $(1-\epsilon)$ the short-run players play best responses to γ_1^* in all but $K(\epsilon)$ periods. The key to our paper is finding an upper bound on this $K(\epsilon)$ that holds uniformly over all equilibria and all discount factors. To do this we view the likelihood ratio corresponding to the short-run player's beliefs about the long-run player's type as a positive supermartingale. Excluding periods where the short-run players play a best response to σ_1^*, this supermartingale is "active" in the sense that in each period where the martingale's value is positive, here is a non-negligible probability of a non-negligible jump. Using theorems about uniform bounds on upcrossing numbers for martingales, we derive uniform bounds on the rate that active supermartingales converge to zero.

To allow the long-run player to build a reputation for playing a mixed strategy, we initially assume there is a positive prior probability that the long-run player is a type who will always use that mixed strategy. Since there are a continuum of mixed strategies for the stage game, in the context of reputations for mixed strategies it may seem more natural to consider models with a continuum of commitment types and a (continuous) prior distribution, so that any particular commitment type has prior probability zero. In the concluding section of the paper we show how our results extend to this case.

2. THE MODEL

The long-run player, player 1, faces an infinite sequence of different short-lived player 2's. Each period, starting with period 0, player 1 selects an action from his action set A_1, while that period's player 2 selects an action from A_2. We assume that players 1 and 2 move simultaneously in each period and that the A_i are finite sets; our earlier paper provided extensions of both of these assumptions. However, in that paper we assumed that the short-run players observed player 1's choice of actions, and we restricted attention to reputations for playing pure strategies. In this paper we will assume that the short-run player's payoffs depend not on player 1's choice of action a_1, but rather on a stochastic "outcome" y_1 which is drawn from a finite set Y_1 with distribution $\rho(y_1|a_1)$. Corresponding to the action spaces A_i are the spaces Σ_i of mixed strategies; when player 1's mixed action is σ_1 the resulting distribution on y_1 is

$$\sum_{{}_1 \in A_1} \sigma_1(a_1) \cdot \rho(y_1|a_1).$$

(Note that this formulation includes the special case where A and Y are isomorphic.) We denote the distribution over outcomes corresponding to strategy σ_1 by $\gamma_1 = \rho \circ \sigma_1$. Since it is unimportant whether or not the short-run players' actions are observable, for simplicity we will assume they are, and identify the space A_2 with a space Y_2 of outcomes of player 2's play.

The short-run players all have the same expected utility function

$$u_2: \quad Y_1 \times Y_2 \to \mathbb{R}.$$

In an abuse of notation, we let $u_2(\sigma) = u_2(\sigma_1, \sigma_2)$ denote the expected payoff corresponding to the mixed strategy $\sigma \in \Sigma$. Each period's short-run player acts to maximize that period's payoff.

Both players know the short-run player's payoff function. On the other hand, player 1 knows his own payoff function, but the short-run players do not. We represent their uncertainty about player 1's payoffs using Harsanyi's [1967] notion of a game of incomplete information. Player 1's payoff is identified with this "type" $\omega \in \Omega$, where Ω is a countable set. It is common knowledge that the short-run players have (identical) prior beliefs μ about ω, represented by a probability measure on Ω.

Let $H = (Y)^{\infty}$ be the measure space of all infinite histories of outcomes, and let $\mathcal{H}$ be the corresponding space of probability measures on H. Player 1's payoff $u_1(h,\omega)$ as depends on the distribution h and his type ω. In particular, for some ω, u_1 may not be additively separable over time, and need not be an expected utility function.

Both long-run and short-run players can observe and condition their play at time t on the entire past history of the realized <u>outcomes</u> of both players, but not on their choice of mixed strategy. (In the case where $Y_1 \cong A_1$, the realized outcome will reveal player 1's choice of action, but not his choice of mixed strategy.) If H_t denotes the set of possible histories (sequences of outcomes) through time t, then a strategy for the period-t player 2 is a map σ_2^t: $H_{t-1} \to \Sigma_2$. Since player 1 knows his type, a strategy for player 1 is a sequence of maps σ_1^t: $H_{t-1} \times \Omega \to \Sigma_1$, specifying his play as a function of history and his type.

We denote this game $G(\delta,\mu)$ to emphasize that it depends on the long-run player's discount factor and on the beliefs of the short-run players.

3. <u>THE THEOREM</u>

Let B: $\Sigma_1 \to \Sigma_2$ be the correspondence that maps mixed strategies by player 1 to the best responses of player 2 (using the payoff u_2). Because the short-run players play only once, in any equilibrium of $G(\delta,\mu)$, each period's play by the short-run player must lie in the graph of B. The short-run

players' behavior can also be characterized by how they respond to distributions over outcomes. Letting Γ_1 be the space of probability distributions over Y_1, we denote this correspondence by β: $\Gamma_1 \to \Sigma_2$. For each strategy σ_1 let $\omega(\sigma_1)$ be the "commitment type" which has "play σ_1 forever" as its strictly dominant strategy for the repeated game, and let $P_1(\Omega,\mu) = \{\omega\in\Omega | \ \omega=\omega(\sigma_1)$ for some σ_1 and $\mu(\omega) > 0\}$ be the set of commitment types which have positive prior probability. In this section we assume that the set of types Ω is countable and that the set P_1 is non-empty. Given that the set of strategies Σ_1 is uncountable, it might be more appealing to consider a <u>density</u> over the set of commitment strategies, so that no single commitment type has positive prior probability. We consider this extension below.

Now fix a type ω_0 whose preferences correspond to the expected discounted value of per-period payoffs:

$$u_1(h,\omega_0) = (1-\delta)E[\sum_{t=0}^{\infty} \delta^t v_1(a_1(t),a_2(t))|(\sigma_1,\sigma_2)]$$

where v_1: $A_1 \times A_2 \to \mathbb{R}$. Given the set of commitment types P_1, which corresponds to reputations the long-run player might be able to maintain, we ask which reputation would be most desirable. Define

$$(1) \qquad v_1^*(P_1) = \max_{\sigma_1\in P_1} \ \min_{\sigma_2 \in B(\sigma_1)} v_1(\sigma_1,\sigma_2),$$

and let $\sigma_1^*(P_1)$ satisfy

$$\min_{\sigma_2\in B(\sigma_1^*)} v_1(\sigma_1^*,\sigma_2^*) = v_1^*.$$

We call $v_1^*(P_1)$ the type-ω_0 commitment payoff relative to the set P_1. Since

we will hold P_1 fixed throughout the paper, we will simplify this to v_1^*; the dependence of what follows on P_1 should be clear. Let $\sigma_1^* = \sigma_1^*(P_1)$ denote the type-ω_0 commitment action, and $\gamma_1^* = \rho \circ \sigma_1^*$ the commitment distribution. (Note that there may be several commitment actions.) Finally, let $\omega^*(=\omega^*(P_1))$ be a type such that such a player 1's best strategy in the repeated game is to play $\sigma_1^*(P_1)$ in every period. Nash equilibrium requires that if ω^* has positive probability, then $\sigma_1^{t-1}(h_t,\omega^*) = \sigma_1^*$ for all t and almost all h_t. We will say that type ω^* is "the" commitment type. Our goal is to argue that with the "right" kind of incomplete-information, type ω_0's worst Nash equilibrium payoff is close to v_1^* when δ is close to one.

Since the game has countably many types and periods, and finitely many actions per type and period, the set of Nash equilibria is a closed non-empty set. This follows from the standard results on the existence of mixed strategy equilibria in finite games, and the limiting results of Fudenberg and Levine [1983, 1986]. Consequently, if $\mu(\omega_0) > 0$, we may define $\underline{V}_0(\delta,\mu)$ to be ω_0 player 1's least payoff in any Nash equilibrium of the game $G(\delta,\mu)$.

Theorem 1: Assume $\mu(\omega_0) > 0$, and that $\mu(\omega^*) \equiv \mu^* > 0$. Then for all $\alpha > 0$, there is a $\underline{\delta} < 1$ such that for all $\delta \in (\underline{\delta},1)$

$$\underline{V}_0(\delta,\mu) \geq (1-\alpha)\, v_1^* + \alpha \min v_1,$$

where $\min v_1$ is the minimum over $A_1 \times A_2$. This says that if type ω_0 is patient relative to the prior probability μ^* that he is "tough", then he can achieve almost his commitment payoff. Moreover, the lower bound on type ω_0's player's payoff is independent of the preferences of the other types in Ω to which μ assigns positive probability. The condition $\mu(\omega_0) > 0$ is necessary only for $\underline{V}_0$ to be well defined.

Proof: We fix an equilibrium $(\hat{\sigma}_1^t, \hat{\sigma}_2^t)$ of $G(\delta,\mu)$, and consider the strategy for player 1 of always playing σ_1^*. The next section of the paper is devoted to showing that for any $\epsilon > 0$, there exists a $K(\mu^*,\alpha)$ otherwise independent of δ and μ, such that player 2's equilibrium strategy chooses actions outside of $B(\sigma_1^*)$ more than $K(\mu^*,\epsilon)$ times with probability no more than ϵ. If we choose $\epsilon = \alpha/2$ and δ sufficiently large that $\delta^{K(\mu^*,\epsilon)} \geq (1-\alpha)/(1-\epsilon)$, then (since $\mu^* \geq 0$) type ω_0 gets at least $(1-\alpha)v_1^* + \alpha \min v_1$. Consequently he gets at least this much in equilibrium. ∎

4. BAYESIAN INFERENCE, SHORT-RUN BEST RESPONSES, AND ACTIVE SUPERMARTINGALES

This section shows that if player 1 plays strategy σ_1^* in every period, it is very likely that the player 2's will choose actions in $B(\sigma_1^*)$ in all but a small number of periods. The key to the proof is a strengthening of the fact that, when player 1 plays strategy σ_1^*, the likelihood ratio corresponding to player 1 not being type ω^* is a positive supermartingale. We strengthen this by observing that in the periods when player 2 does not play a best response to σ_1^*, the odds ratio is an "active supermartingale" in the sense of having a non-negligible probability of jumping a non-negligible amount. While positive supermartingales can converge to positive limits, or decrease towards zero very slowly, the Appendix shows that positive supermartingales which are active in our sense converge to zero at a uniform rate that depends only on their inital value and their degree of "activity."

Let $h_t \in H_t$ be identified with the subset of histories $h \in H$ that coincide with h_t through and including period t. In this way H_t may be viewed as a subset of H. Type ω_0 wishes to calculate his payoffs if he plays σ_1^* in every period, given the equilibrium $(\hat{\sigma}_1^t,\hat{\sigma}_2^t)$. Thus he should use the measure over H defined by $(\sigma_1^*,\hat{\sigma}_2^t)$. Subsequently, we use this measure and H as our underlying sample space. However, the player 2's use player 1's equilibrium strategy $\hat{\sigma}_1^t(h_t,\omega)$ in computing their beliefs and optimal responses. Let

$\mu(\omega|h_{t-1})$ be the conditional probability distribution over player 1's types obtained by updating the prior $\mu(\omega)$ in accordance with Bayes law and the equilibrium strategy $\hat{\sigma}_1$.

Number the outcomes in Y_1 from 1 to n, and let p_k be the probability that the commitment distribution $\gamma_1^* = \rho \circ \sigma_1^*$ assigns to outcome k. Let $q(h_{t-1})$ be the distribution over time-t outcomes predicted by $\hat{\sigma}$ conditional on the history being h_{t-1} and $\omega \in \Omega_0 \equiv \Omega/\omega^*$:

$$q(h_{t-1}) = \sum_{\omega\in\Omega_0} \mu(\omega|h_{t-1})\ \rho\circ\hat{\sigma}_1^t(h_{t-1},\omega))\ /\ (1-\mu(\omega^*|h_{t-1})).$$

Let $d(\gamma_1,\gamma_1') = \sup_k \|\gamma_1(y_k)-\gamma_1'(y_k)\|$ be the distance in the sup norm between distributions γ_1 and γ_1', and set $\Delta(h_{t-1}) = d(q(h_{t-1}),\gamma_1^*)$. Also, define $\nu(h_{t-1}) = (1-\mu(\omega^*|h_{t-1}))q(h_{t-1})+ \mu(\omega^*|h_{t-1})p$. This is the probability distribution over outcomes that player 2 expects to face in period t. Our first claim is that if the equilibrium distribution $q(h_{t-1})$ is close to the commitment distribution p in the sense that Δ is small, then ω_0 gets at least v_1^* in period t.

<u>Lemma 1</u>: There is a number $\Delta_0>0$ such that if h_{t-1} has positive probability and $\Delta(h_{t-1}) \le \Delta_0$, then the equilibrium strategy of player 2 gives type ω_0 of player 1 at least v_1^* at time t.

<u>Proof</u>: Observe that $d(\nu(h_{t-1}),\ \gamma_1^*) < \Delta(h_{t-1})$, and that v_1^* is defined relative to the best response to γ_1^* that type ω_0 likes the least. Since player 2's best response correspondence B is upper hemi-continuous, for ν close to γ_1^* we know that each element of $B(\nu)$ must be close to an element of $B(\gamma_1^*)$. Now since player 2 has a finite number of pure strategies, a strategy σ_2 can be close to an element of $B(\gamma_1^*)$ only if it places probability close to one on

pure strategies in the support of $B(\gamma_1^*)$. And since player two must be indifferent between all strategies he is willing to assign positive probability, we conclude that the support of $B(\nu)$ must be contained in the support of $B(\gamma_1^*)$ for ν sufficiently close to γ_1^*. The conclusion of the lemma follows immediately. ∎

Define families of random variables $(\hat{p}_t(h), \hat{q}_t(h))$ by setting $\hat{p}_t = p_k$ and $q_t = q_k(h_{t-1})$ if y_1^k occurs at time t. Define another family of random variables $L_t(h)$ as follows: For t=0 set

$$L_0(h) \equiv \frac{1-\mu(\omega^*)}{\mu(\omega^*)} .$$

Then let $h_t \in H_t$ be the finite history that coincides with h through and including time t and define recursively

$$L_t(h) \equiv \frac{\hat{q}_t(h)}{\hat{p}_t(h)} L_{t-1}(h) .$$

It is well known that $L_t(h) = [1-\mu(\omega^*|h_t)]/\mu(\omega^*|h_t)$ is player 2's posterior odds ratio that player 1 is not type ω^*. It is also well known that this odds ratio is a supermartingale. We give a proof for completeness:

<u>Lemma 2</u>: $L_t(h) = [1-\mu(\omega^*|h_t)]/\mu(\omega^*|h_t)$ and (L_t, H_t) is a supermartingale.

<u>Proof</u> : The first claim is true for L_0. Imagine it is true for L_{t-1}, then

$$[1-\mu(\omega^*|h_t)]/\mu(\omega^*|h_t) = \hat{q}_t(h)[1-\mu(\omega^*|h_{t-1})]/[\hat{p}_t\mu(\omega^*|h_{t-1})]$$

$$= (\hat{q}_t/\hat{p}_t)\ L_{t-1} = L_t .$$

To see that L_t is a supermartingale, observe that

$$E[L_t | L_{t-1}, h_{t-1}] = L_{t-1} \sum_{k \in supp(p)} p_k [q_k(h_{t-1})/p_k]$$

$$= L_{t-1} \sum_{k \in supp(p)} q_k(h_{t-1}) \leq L_{t-1}. \qquad \blacksquare$$

Lemma 3: If h_{t-1} has positive probability and $L_{t-1} \leq \Delta_0$ then the equilibrium strategy of player 2 gives type ω_0 at least v_1^* (with probability one) at h_t.

Proof: If $L_{t-1} \leq \Delta_0$ then $1-\mu(\omega^*|h_{t-1}) \leq \mu(\omega^*|h_{t-1})\Delta_0 \leq \Delta_0$. Since $d(\nu(h_{t-1}),p) \leq 1-\mu(\omega^*|h_{t-1}) \leq \Delta_0$, the conclusion follows from Lemma 1. $\blacksquare$

Lemma 4: If $\Delta(h_{t-1}) > \Delta_0$ then $Pr[L_t/L_{t-1} - 1 \leq -\Delta_0/n \mid h_{t-1}] \geq \Delta_0/n$ almost surely.

Proof: Note first that $L_t/L_{t-1} = \hat{q}_t/\hat{p}_t$, which is $q_1(h_{t-1})/p_1$ with probability p_1; $q_2(h_{t-1})/p_2$ with probability p_2, and so forth for those indices k for which $p_k \neq 0$. Consequently, it suffices to show for some k

$$q_k(h_{t-1})/p_k \leq 1 - \Delta_0/n \text{ and } p_k \geq \Delta_0/n.$$

Suppose, without loss of generality, that $\Delta(h_{t-1}) = \| p_1 - q_1(h_{t-1}) \| \geq \Delta_0$. If $p_1 - q_1(h_{t-1}) \geq \Delta_0$ then $p_1 \geq \Delta_0$, and $1 - q_1(h_{t-1})/p_1 \geq \Delta_0/p_1 \geq \Delta_0$, and we are done. If, on the other hand, $q_1(h_{t-1}) - p_1 \geq \Delta_0$, then $\Sigma_{k>1} (p_k - q_k(h_{t-1})) \geq \Delta_0$. Consequently $n \max_{k>1} (p_k - q_k(h_{t-1}) \geq \Delta_0$, and, for k=2 say, we have $p_2 - q_2(h_{t-1}) \geq \Delta_0/n$. Again, we conclude $p_2 \geq \Delta_0/n$ and $1 - q_2(h_{t-1})/p_2 \geq \Delta_0/n$. $\blacksquare$

Lemma 4 shows that in the periods where the marginal distribution on the actions of the non-commitment types differs significantly from the commitment distribution, the likelihood ratio is likely to jump down by a significant amount. Of course, in periods where $\Delta(h_{t-1})$ is small, the likelihood ratio need not change much, but in these periods we know from Lemma 1 that player two will play a best response to the commitment strategy, and so the payoff of type ω_0 of player one is at least v_1^*. The key to our result is to show that with high probability there are few periods where player one's payoff is less than v_1^*, that is, few periods where $\Delta(h'_{t-1}) > \Delta_0$. We will call these bad periods. To show that there are unlikely to be many bad periods, we introduce a new supermartingale which includes all of the bad periods from the supermartingale L.

We first define a sequence of stopping times. Set $\tau_0 = 0$. If $\tau_{k-1}(h) = \infty$, set $\tau_k(h) = \infty$ as well. If $\tau_{k-1}(h)$ is finite, set $\tau_k(h)$ to be the first time $t > \tau_{k-1}(h)$ such that either

(1) $\quad P_r [\| L_t/L_{t-1} - 1 \| < \Delta_0/n] \geq \Delta_0/n$, or

(2) $\quad L_t/L_{\tau_{k-1}} - 1 \geq \Delta_0/2n$, or

(3) $\quad$ if no such time exists, set $\tau_k(h) = \infty$.

Lemma 4 shows that this sequence of stopping times picks out all the bad date-history pairs, that is, those for which $\Delta(h_{t-1}) > \Delta_0$.

The <u>faster process</u> $\tilde{L}_k$ is defined by $\tilde{L}_k = L_{\tau_k}$. Since the τ_k are stopping times, $\tilde{L}_k$ is a supermartingale, with an associated filtration whose events we denote h_k. Moreover, we will show that $\tilde{L}_k$ is an "active" supermartingale in the following sense:

Definition: A process $\tilde{L}_t$ with nonnegative values is an active supermartingale with activity ψ if

$$\Pr[\| \tilde{L}_{t+1}/\tilde{L}_t - 1 \| > \psi \,|h_t) > \psi$$

for all histories h_t such that $\tilde{L}_t > 0$.

Lemma 5: $\tilde{L}_k$ is an active supermartingale with activity $\Delta_0/2n$.

Proof: Since the τ_k are stopping times, $\tilde{L}_k$ is a supermartingale. Next, we claim that if h is such that $\tilde{L}_{k-1} > 0$,

$$\Pr[\| \tilde{L}_k/\tilde{L}_{k-1} - 1\| > \Delta_0/2n \mid h_{k-1}] > \Delta_0/n.$$

To see this, let $s = \tau_{k-1}(h)$, which is a constant with respect to h_{k-1}; $\tau_k(h_{k-1})$ is a random variable. We will show that

$$\Pr[\|L_{\tau_k}/L_s - 1\| > \Delta_0/2n \mid h_s] > \Delta_0/n.$$

One of the three rules in the definition of the τ's must be used to choose τ_k. We will show that this inequality holds conditional on each rule, and thus that it holds averaging over all of them. Conditional on h_s, if rule (1) or (3) is used, then with probability one $\| L_{\tau_k}/L_s - 1 \| > \Delta_0/2n$. If rule 2 is used,

$$\Pr[L_{\tau_k}/L_{(\tau_k-1)} - 1 < -\Delta_0/n \mid h_s, \{\text{rule 2 used}\}] > \Delta_0/n,$$

and also since rule 1 was not used at $\tau_k - 1$,

$$L_{(\tau_k-2)}/L_s - 1 < \Delta_0/2n.$$

Combining the last two inequalities shows that

$$\Pr[L_{\tau_k}/L_s - 1 < (-\Delta_0/n + \Delta_0/2n - \Delta_0^2/2n^2) \mid h_s, \text{(rule 2)}] > \Delta_0/n.$$

Since $(-\Delta_0/n + \Delta_0/2n - \Delta_0^2/2n^2) < -\Delta_0/2n$, we conclude that

$$\Pr[\|L_{\tau_k}/L_s - 1\| > \Delta_0/2n \mid h_s, \text{(rule 2)}] > \Delta_0/n. \qquad \blacksquare$$

Next we state a key part of the proof: active supermartingales converge to zero at a uniform rate that depends only on their initial value and their degree of activity.

<u>Theorem A.1</u>: Let $\ell_0 > 0$, $\psi \in (0,1)$, and $\epsilon > 0$ be given. For each $0 < \underline{L} < \ell_0$, and each $\epsilon > 0$, there is a time $K < \infty$ such that

$$\Pr[\sup_{k>K} \tilde{L}_k \le \underline{L}] \ge 1 - \epsilon$$

for every active supermartingale $\tilde{L}$ with $\tilde{L}_0 = \ell_0$ and activity ψ.

This theorem is proved in the Appendix using results about upcrossing numbers. The key aspect of the Theorem is that the bound K depends only on ℓ_0 and ψ, and is independent of the particular supermartingale chosen.

To conclude, we recapitulate the proof of Theorem 1. From Lemmas 3 and 4 we know that for all histories where player 1 has always played σ_1^*, type ω_0 receives at least v_1^* in all periods t except possibly those where $t = \tau_k(h)$ for some k. If we set $\tilde{L}_0 = (1-\mu^*)/\mu^*$ and $\underline{L} = (1-\Delta_0)/\Delta_0$ in Theorem A.1, we see that for all $\epsilon > 0$ there is a $K(\mu^*,\epsilon)$ such that with probability at least $(1-\epsilon)$ type ω_0 receives v_1^* in all but $K(\mu^*,\epsilon)$ periods. This implies that

$$\underline{V}_0(\delta,(\Omega,\mu)) \ge (1-\epsilon)\delta^{K(\mu^*,\epsilon)} v_1^* + [1-(1-\epsilon)\delta^{K(\mu^*,\epsilon)}] \min v_1.$$

Thus for any $\alpha > \epsilon$, by setting δ large enough that $\delta^{K(\mu^*,\epsilon)} > (1-\alpha)/(1-\epsilon)$ we have that $\underline{V}_0(\delta,\mu) \ge (1-\alpha)v_1^* + \alpha \min v_1$.

5. CONTINUUM OF COMMITMENT TYPES

The results above treat the case where the set Ω of commitment types is countable. In this Section we consider the case where Ω includes a single "sane" type ω_0 which has positive prior probability, i.e. $\mu(\omega_0) > 0$, and a continuum of commitment types" $\omega(\sigma_1)$ corresponding to each of player 1's mixed strategies $\sigma_1 \in \Sigma_1$, (and possibly other types as well). The probability distribution over commitment types is given by a continuous density $d\mu(\omega)$.

For each strategy σ_1, define the sane type 's corresponding commitment payoff:

$$v_1^*(\sigma_1) = \min_{\sigma_2 \in B(\sigma_1)} v_1(\sigma_1, \sigma_2; \omega_0).$$

We will prove that type ω_0 can approximate the commitment payoff to any σ_1 if the discount factor δ is sufficiently close to one.

<u>Theorem 2</u> : Assume $\mu(\omega_0) > 0$, and that $d\mu$ is uniformly bounded below by $\eta>0$ over all of the commtiment types $\omega(\sigma_1)$. Then for all σ_1, and all $\alpha > 0$, there is a $\underline{\delta} < 1$ such that for all $\delta \in (\underline{\delta}, 1)$, type ω_0's payoff is at least

$$(1-\alpha)\, v_1^*(\sigma_1) + \alpha \min v_1$$

in any Nash equilibrium of $G(\delta, \mu)$.

To prove the theorem, fix a Nash equilibrium $(\hat{\sigma}_1^t, \hat{\sigma}_2^t)$, and a $\sigma_1 \in \Sigma_1$ with $\gamma_1 = \rho \circ \sigma_1$. For each $\epsilon>0$, let N_ϵ be the ϵ-neighborhood of σ_1 in the supremeum norm. As before, let $\nu(h_{t-1})$ be the distribution over outcomes implied by $(\hat{\sigma}_1, \hat{\sigma}_2)$ when the history is h_{t-1}, let $q(h_{t-1})$ be the probability distribution over outcomes at h_{t-1} conditional on ω being in the complement of N_ϵ.

Lemma 6 : There is an $\underline{\epsilon} > 0$ such that if $\epsilon<\underline{\epsilon}$ then $B(\sigma_1') \subsetneq B(\sigma_1)$ for all $\sigma_1' \in N_\epsilon$. Fixing $\epsilon<\underline{\epsilon}$, there is a $\Delta_0>0$ such that if h_{t-1} has positive probability and $d(q(h_{t-1}), \gamma_1)<\Delta_0$ then player 2 will play a best response to σ_1 at time t.

Proof: Essentially the same as Lemma 1: the keys are the upper hemicontinuity of B and the assumption that player two has only finitely many pure strategies. ∎

For each history h_{t-1} with positive probability, let $d\mu(\omega(\sigma_1)|h_{t-1})$ be the conditonal distribution over commitment types derived from the equilibrium strategies. In the proof of Theorem 1, we considered the strategy for type ω_0 of always playing a fixed mixed strategy σ_1^*. In the present case it will be more convenient to consider a slightly more complicated strategy for type ω_0. Specifically, define a history-dependent sequence of distributions p on the outcome space Y_1 by

$$p(h_{t-1}) = \int_{\sigma_1 \in N_\epsilon} (\rho \circ \sigma_1)\, d\mu(\omega(\sigma_1)|h_{t-1}) \Big/ \int_{\sigma_1 \in N_\epsilon} d\mu(\omega(\sigma_1)|h_{t-1}) .$$

Define a family of random variables $(\hat{p}_t(h), \hat{q}_t(h))$ by setting $\hat{p}_t = p_k(h_{t-1})$ and $q_t = q_k(h_{t-1})$ if y_1^k occurs at time t, and define a second family L_t as follows: In period 0,

$$L_0(h) = \left[1 - \int_{\sigma_1 \in N} d\mu(\omega(\sigma_1))\right] \Big/ \int_{\sigma_1 \in N} d\mu(\omega(\sigma_1))$$

Then define recursively

$$L_t(h) = \frac{\hat{q}_t(h)}{p_t(h)} L_{t-1}(h).$$

This is the likelihood ratio for player 1 not being a type in N_ϵ. The key change required to our earlier proof is that if player 1 adopted the strategy of always playing σ_1 the likelihood ratio would not be a supermartingale, as its behavior at each date would depend on the relative weights given to types ω in N_ϵ. However, if player 1 adopts the strategy $\tilde{\sigma}_1^t$ defined by

$$\tilde{\sigma}_1^t = \int_{\sigma_1 \in N_\epsilon} \sigma_1 \, d\mu(\omega(\sigma_1))|h_{t-1}) \Big/ \int_{\sigma_1 \in N_\epsilon} d\mu(\omega(\sigma_1))|h_{t-1}),$$

that is, if he plays to mimic the average expected play of types in N_ϵ, then it is easy to see that L_t is a supermartingale.

Lemma 7 : If player 1 adopts the strategy of playing $\tilde{\sigma}_1^t(h_{t-1})$ at each time t, then (L_t, h_t) is a supermartingale and

$$L_t(h) = \left[1 - \int_{\sigma_1 \in N_\epsilon} d\mu(\omega(\sigma_1))|h_{t-1}) \right] \Big/ \left[\int_{\sigma_1 \in N_\epsilon} d\mu(\omega(\sigma_1))|h_{t-1}) \right].$$

Proof : Same as Lemma 2. ■

From here on we can follow the proofs of Lemmas 3 through 5, replacing the event $\omega = \omega^*$ with $\omega \in N_\epsilon$ and replacing the stragegy σ_1^* with $\tilde{\sigma}_1^t$. We then extract the faster process $\tilde{L}$, which picks out the "bad" periods according to conditions 1 through 3 on page 15, and apply Theorem A.1 to conclude that there are unlikely to be many bad periods. Since player 1 receives at least $v_1^*(\sigma_1)$ in the good periods by Lemma 6, the conclusion of Theorem 2 follows.

REFERENCES

Aumann, R. and S. Sorin [1987], "Cooperation and Bounded Rationality."

Fudenberg, D. and D. Kreps [1987], "Reputation and Simultaneous Opponents," _Review of Economic Studies_ 54, 541-568.

Fudenberg, D. D. Kreps and D. Levine [1987], "On the Robustness of Equilibrium Refinements," forthcoming in the _Journal of Economic Theory_.

Fudenberg, D., D. Kreps and E. Maskin [1987], "Repeated Games with Long- and Short-Lived Players," MIT Working Paper No. 474.

Fudenberg, D. and D. Levine [1983], "Subgame-Perfect Equilibria of Finite and Infinite Horizon Games", _Journal of Economic Theory_, 31, 251-268.

__________ [1986], "Limit Games and Limit Equilibria," _Journal of Economic Theory_, 30, 261-279.

__________ [1988], "Reputation and Equilibrium Selection in Games with a Single Patient Player," MIT Working Paper No. 461.

Fudenberg, D. and E. Maskin [1986], "The Folk Theorem in Repeated Games with Discounting or with Incomplete Information," _Econometrica_, 54, 533-554.

Harsanyi, J. [1967], "Games with Incomplete Information Played by Bayesian Players," _Management Science_, 14, 159-82, 320-34.

Kreps, D., P. Milgrom, J. Roberts, and C. Wilson [1982], "Rational Cooperation in the Finitely Repeated Prisoners' Dilemma," _Journal of Economic Theory_, 27, 245-252, 486-502.

Kreps, D. and C. Wilson [1982], "Reputation and Imperfect Information," _Journal of Economic Theory_, 27, 253-279.

Loève, M. [1978], _Probability Theory II_, Springer-Verlag, New York.

Milgrom, P. and J. Roberts [1982], "Predation, Reputation and Entry Deterrence," _Econometrica_, 50, 443-460.

Neveu [1975], _Discrete-Parameter Martingales_, North Holland, Amsterdam.

Selten, R. [1977], "The Chain-Store Paradox," _Theory and Decision_, 9, 127-159.

Our goal is to prove

Theorem A.1: Let $\ell_0 > 0$, $\epsilon > 0$, and $\psi \in (0,1)$ be given. For each $\underline{L}$, $0 < \underline{L} < \ell_0$, there is a time $K < \infty$ such that $\Pr(\sup_{k>K} \tilde{L}_k \leq \underline{L}) \geq 1-\epsilon$ for every active supermartingale $\tilde{L}$ with $\tilde{L}_0 = \ell_0$ and activity ψ. ■

For a given martingale the above is a simple consequence of the fact that $\tilde{L}$ converges to zero with probability one. The force of the theorem is to give a uniform bound on the rate of convergence for all supermartingales with a given activity ψ and initial value ℓ_0.

Throughout the appendix, we use $\tilde{L}$ to denote any supermartingale that satisfies the hypotheses of Theorem A.1. To prove the theorem, we will use some fundamental results from the theory of supermartingales, in particular bounds on the "upcrossing numbers" which we introduce below. These results can be found in Neveu [1975], Chapter II.

Fact A.2: For any supermartingale, $\Pr[\sup_{k>K} \tilde{L}_k \geq c] \leq \min(1, \tilde{L}_0/c)$.

Next, fix an interval $[a,b]$, $0 < a < b < \infty$, and define $U_k(a,b)$ to be the number of "upcrossings" of $[a,b]$ up to time k; let $U_\infty(a,b)$ be the total number of upcrossings (possibly equal to ∞).

Fact A.3: $\Pr[U_\infty(a,b) \geq N] \leq (a/b)^N \min(\tilde{L}_0/a, 1)$.

This is known as Dubin's inequality. (See, for example, Neveu [1975], p. 27).

Next we observe that since $\tilde{L}$ has activity ψ, it makes a jump of size ψ with probability at least ψ in each period k where $\tilde{L}_k$ is nonzero. Consequently, over a large number of periods either $\tilde{L}$ has jumped to zero or

there are likely to be "many" jumps. Specifically, define J_k to be the number of times $k' < k$ that $\|\tilde{L}_{k'+1}/\tilde{L}_{k'} - 1\| > \psi$.

Lemma A.4: For all ϵ and J there exists a K such that

$\Pr[(J_K \geq J) \text{ or } (\tilde{L}_K = 0)] \geq \psi.$

Proof: Because $\tilde{L}$ has activity ψ, in each period k', either $\tilde{L}_{k'} = 0$ or the probability of a jump of size ψ at time k' exceeds ψ. Define a sequence of indicator functions I_k by $I_k = 1$ iff $(\tilde{L}_k = 0$ or $\|\tilde{L}_k/\tilde{L}_{k-1} - 1\| > \psi)$, and set $S_K = \sum_{k<K} I_k$. Each I_k has expectation at least ψ, so for some K sufficiently large, $\text{Prob}[S_K \geq J] \geq 1-\epsilon$. Now if $S_K \geq J$, then either $\tilde{L}_K = 0$ for some $k \leq K$, in which case $\tilde{L}_k = 0$ as well, or there have been at least J jumps by time K. ∎

We have now established that most paths of $\tilde{L}$

(1) Do not exceed $\bar{c}$ for $\bar{c}$ large, (Fact A.2)

(2) Make "few" upcrossings of any positive interval [a,b] (Fact A.3), and

(3) Either make "lots of jumps" or hit zero. (Lemma A.4)

We will use these three conditions to show that for K large, most paths remain below $\underline{L}$ from K on. To do so, we first argue that most paths will pass below $\underline{c}$ by time K.

Divide the interval $[\underline{c}, \bar{c}]$ into I equal subintervals with endpoints $e_1 = \underline{c}, \ldots, e_{I+1} = \bar{c}$. Then define the events

E_1 if $\max_{k \leq K} \tilde{L}_k \geq \bar{c}$;

E_2 if at least one of the intervals $[e_i, e_{i+1}]$ is upcrossed N or more times;

E_3 if $J_K < J$ and $\tilde{L}_K > 0$,

E_4 if $\min_{k\le K} \tilde{L}_k < \underline{c}$.

By judicious choice of $\bar{c}$, I, K, N and J, we will insure that $E_4 \subset E_1^c \cup E_2^c \cup E_3^c$ and that $\Pr(E_1)$, $\Pr(E_2)$, $\Pr(E_3) \le \epsilon/3$. This will yield our preliminary conclusion that

$$\Pr[\min_{k\le K} \tilde{L}_k < \underline{c}] = \Pr(E_4) \le 1-\epsilon.$$

If we choose

$$\underline{c} = (\epsilon/3)\ell_0$$

Fact A.2 implies that

$$\Pr[\max_{k>K} \tilde{L}_k > \underline{L} \mid \min_{k\le K} \tilde{L}_k < \underline{c}]\underline{c}/\underline{L} \le \epsilon/3$$

giving us the desired conclusion that

$$\Pr[\max_{k>K} \tilde{L}_k > \underline{L}] \le (1-\epsilon)\ \epsilon/3.$$

Turning first to E_1, we can again use Fact A.2 to choose

$$\bar{c} = (3/\epsilon)\ell_0$$

and insure that $\Pr(E_1) = \Pr(\max_{k\le K} \tilde{L}_k \ge \bar{c}) \le \epsilon/3$. Note for future reference that this is true, regardless of how we pick K.

In the range above $\underline{c}$, when $\|\tilde{L}_k/\tilde{L}_{k-1} - 1\| > \psi$, $\|\tilde{L}_k - \tilde{L}_{k-1}\| \ge \psi\underline{c}$. Thus, if we choose

$$I \ge 2\bar{c}/\underline{c}\psi + 1$$

and if $\tilde{L}_k \ge (1+\psi)\underline{c}$, then there is at least a ψ chance of crossing one of the subintervals $[e_i, e_{i+1}]$. On the other hand, a path that remains between $\underline{c}$ and $\bar{c}$ and has J or more jumps across subintervals must cross at least one

subinterval $(J-I)/2I - 1$ times. Consequently if we choose

$$(*) \qquad N \leq (J-I)/2I - 1$$

then $E_4 \subset E_1^c \cup E_2^c \cup E_3^c$ as required. In other words, a path that does not go above $\bar{c}$, that does not upcross any subinterval in $[\underline{c},\bar{c}]$ N or more times, and jumps K or more times, must fall below $\underline{c}$. By Fact A.3, we know that for any given subinterval, the probability of N or more upcrossings is not more than

$$(1+\psi)^{-N} \ell_0/\underline{c}.$$

Consequently, the probability that some subinterval is upcrossed N or more times is no more than

$$I(1-\psi)^{-N} \ell_0/\underline{c}.$$

To make $Pr(E_2) \leq \epsilon/3$ we should choose

$$N \geq \frac{3I\ell_0/\underline{c}\epsilon}{\log(1-\psi)}$$

This determines J by (*) above

$$J = 2I(N+1) + I.$$

Finally, choose K by fact A.4 to make $Pr(E_3) \leq \epsilon/3$. ∎

3-22-2'

6-5-89

Date Due

FEB 13 1991		
MAR 01 1999		